

# Long-Term Clinical Outcomes of Care Management for Chronically Depressed Primary Care Patients: A Report From the Depression in Primary Care Project

Michael S. Klinkman, MD, MS

Sabrina Bauroth, MSW

Stacey Fedewa, MPH

Kevin Kerber, MD

Julie Kuebler, CNP

Tanya Adman, MSW

Ananda Sen, PhD

University of Michigan Health System,  
Ann Arbor, Michigan



*Conflicts of interest: Michael S. Klinkman has received related grant support from the National Institute of Mental Health, the Lilly Foundation, and Robert Wood Johnson Foundation, and has served on scientific advisory boards for Wyeth Pharmaceuticals; Sabrina Bauroth has received related grant support from the Lilly Foundation; Stacey Fedewa has no conflicts to report; Kevin Kerber has received related grant support from the Robert Wood Johnson Foundation and Lilly Foundation; Julie Kuebler has received related grant support from the Lilly Foundation; Tanya Adman has no conflicts to report; Ananda Sen has no conflicts to report.*

## CORRESPONDING AUTHOR

Michael Klinkman, MD, MS  
Departments of Family Medicine and  
Psychiatry  
University of Michigan Medical School  
1018 Fuller St  
Ann Arbor, MI 48109-0708  
mklinkma@umich.edu

## ABSTRACT

**PURPOSE** Recent studies examining depression disease management report improvements in short-term outcomes, but less is known about whether improvements are sustainable over time. This study evaluated the sustained clinical effectiveness of low-intensity depression disease management in chronically depressed patients.

**METHODS** The Depression in Primary Care (DPC) intervention was introduced in 5 primary care practices in the University of Michigan Health System, with 5 matched practices selected as control sites. Clinicians were free to refer none, some, or all of their depressed patients at their discretion. Core clinical outcomes of remission and serial change in Patient Health Questionnaire (PHQ-8) scores for 728 DPC enrollees observed for up to 18 months after enrollment were compared with those for 78 patients receiving usual care who completed mailed questionnaires at baseline, 6, 12, and 18 months.

**RESULTS** DPC enrollees had sustained improvement in remission rates and reduced-function days over the full 18 months. Mean change in the PHQ-8 score over each 6-month interval was more favorable for DPC enrollees than for usual care patients, and the proportion of DPC enrollees in remission was higher at 6 months (43.4% vs 33.3%;  $P = .11$ ), 12 months (52.0% vs 33.9%;  $P = .012$ ), and 18 months (49.2% vs 27.3%;  $P = .004$ ). Multivariate analysis controlling for age, sex, ethnicity, baseline severity, and comorbid medical illness confirmed that DPC enrollees had significantly more reduction in depressive symptom burden over 18 months.

**CONCLUSIONS** The DPC intervention produced sustained improvement in clinical outcomes over 18 months in a cohort of chronically depressed patients with persistent symptoms despite active treatment.

*Ann Fam Med* 2010;8:387-396. doi:10.1370/afm.1168.

## INTRODUCTION

Inadequate treatment of depressed patients in primary care settings remains a major public health problem despite decades of study.<sup>1</sup> Efforts to improve care have moved from a narrow focus on increasing detection rates through development and testing of comprehensive clinical support protocols integrated into primary care practices.<sup>2-9</sup> Despite intense efforts to develop and refine these protocols, the longer term effectiveness and sustainability of clinician support protocols has in many cases fallen short of early expectations.<sup>9-15</sup>

The central problem remains that the acute-phase treatment model embedded in these efforts is a poor fit for the chronicity and complexity of the mental health problems seen by primary care clinicians in routine

practice. Depressive disorder is a chronic condition, and in primary care practices most patients with this disorder have chronic or recurrent depression, with many cases further complicated by comorbid mental health disorders.<sup>16-20</sup> Treatment protocols designed to improve the effectiveness of acute-phase care apply to only a small fraction of the depressed patients seen by primary care physicians.<sup>3,9</sup> It is unrealistic to expect that protocols which offer little benefit for the great majority of depressed primary care patients will be widely implemented.

The Depression in Primary Care (DPC) project was 1 of several demonstration projects funded by the Robert Wood Johnson Foundation as part of their Depression in Primary Care: Linking Clinical and Systems Strategies initiative.<sup>15,21,22</sup> The primary aim of the DPC project was to develop, implement, and evaluate the effectiveness and sustainability of a depression management program intended to support primary care clinicians' management of patients in both acute and chronic phases of treatment. The DPC project was developed through a collaborative effort that involved several University of Michigan Health System (UMHS) entities (the University of Michigan Medical Management Center, the Departments of Family Medicine and Psychiatry, the University of Michigan Depression Center, the Medical Center Information Technology unit, and the Faculty Group Practice), along with M-Care HMO (health maintenance organization), the M-WORKS disability management program, and the Ford Motor Company.

In this article, we describe the key design features of the DPC intervention, report the primary clinical outcomes for DPC enrollees, and compare core clinical outcomes for intervention (DPC enrollees) vs control group (usual care) patients over the 18-month demonstration period.

## METHODS

### The DPC Intervention

The DPC clinical intervention was designed to follow the Chronic Care Model<sup>23,24</sup> and included the following components.

#### Care Management

Central care managers were assigned to support care at a specific site. Care managers were either registered nurses or had a master's degree in social work and experience in direct provision of mental health services. The care manager contacted all referred patients by telephone, assessed depression severity and complexity with the assistance of a consultation-liaison psychiatrist, where necessary, and carried out a series

of monitoring and support telephone calls to patients at intervals based on severity and complexity.

#### Disease Monitoring and Clinician Feedback

The outcome-monitoring protocol included the Patient Health Questionnaire-8 (PHQ-8, the scored version of the PHQ-9, with item 9 on self-harm omitted),<sup>25</sup> the SF-12 Health Survey,<sup>26</sup> and items on treatment adherence, functional impairment, and satisfaction with care (the full protocol is available on request). Patients completed this protocol at differing intervals (ranging from 1 to 6 months, depending on clinical severity) through a direct call from the care manager or by a prompted call to an interactive voice response (IVR) system. After each call an outcome summary report, including results, graphed trend lines, and patient-specific evidence-based treatment options, was sent to the referring clinician by e-mail or fax; later in the demonstration project, this report was electronically posted to the electronic medical record used by all participating clinicians.

#### Patient Activation and Self-Management Assistance

An educational curriculum that focused on activation and self-management skills was designed by program staff and administered by care managers, initially through small-group sessions and later by integration into individual monitoring and support calls. All enrollees were sent a short printed manual on self-management and activation after the intake call, and care managers worked with enrollees to set a self-management goal in the next call. Progress toward goals was monitored during follow-up calls.

#### Clinical Information System

A depression disease registry, created from existing UMHS databases and periodically updated, provided the denominator for several population-based analyses. A secure e-mail system enabled care managers, primary care physicians, and the consultation-liaison psychiatrist to communicate efficiently to coordinate care. The UMHS electronic medical record was used by care managers and clinicians to monitor patients' clinical progress. The project team developed a new case management software application to assist care managers in tracking patients enrolled in DPC project. This application also compiled clinical case notes and stored the results of all outcome assessments.

#### Physician Reimbursement

Internal procedure codes were created and assigned a level of relative value unit (RVU) credit. Care managers were asked to record all instances of care coordination with the physicians, with RVUs credited to the physician. Surprisingly, these codes were rarely turned

in for reimbursement during the demonstration period, even though their occurrence was routinely tracked by care managers to spare physicians the burden of reporting. There were 2 main reasons: first, the level of RVU credit would in most cases have amounted to only a small direct clinical incentive payment, projected at less than \$300 for a physician per year at a typical rate of coordination activities; presumably, this figure was too low to induce physicians to print and sign an attestation form to receive RVU credit. Second, in qualitative interviews, several physicians stated that the program added value to their practice, and they were reluctant to ask for additional payment.

### Integration of Disability Management

An expedited link to the UMHS disability management program was created to enable primary care physicians to certify work disability from depression when coupled to tailored disability management.

### Content of DPC Intervention

The DPC intervention consisted of a series of telephone calls and e-mail exchanges between enrollee and care manager, care manager and referring physician, and, on occasion, care manager and consultation-liaison psychiatrist. Most contacts between care manager and enrollees were by telephone. The 2 primary goals of care management were to increase enrollees' self-management of depression and provide feedback to the referring physician about clinical progress and possible complications in treatment. Before the formal start date, care managers and project staff created resource manuals that included social and community service agencies, specialty mental health service providers, and local key contacts for each of the 3 geographic areas covered by the intervention.

The first (intake) telephone call, which required 20 to 30 minutes, comprised a review of the enrollee's clinical history, introduction to the program components, and a first discussion of depression self-management and goal-setting; the patient was then transferred to the IVR system to complete the set of outcome measures. Upon completion patients were mailed a packet of written materials on depression, self-management, and the DPC program (available from author on request). The second telephone call, which usually lasted less than 10 minutes, focused on answering questions from the materials and setting a self-management goal. Subsequent calls focused on support, monitoring progress toward goals, and completing IVR-based outcome measures. Specific support needs were addressed as they emerged (for example, referral to local social services or community agencies). Guidelines for scheduling follow-up telephone calls based upon level of complexity were

used, but care managers were also encouraged to use their clinical discretion when scheduling contacts with patients. As a rule, call frequency declined as enrollees improved, with maintenance calls scheduled at 6-month intervals. Figure 1 illustrates this process with a case example.

Care managers were trained to use motivational interviewing techniques when working with patients on self-management, but to not apply the specific therapy techniques used in telephone-based interpersonal therapy or cognitive behavioral therapy. During training, calls were monitored and reviewed to minimize provision of advice or guidance that approximated interpersonal or cognitive behavioral therapy. During ongoing supervision, telephone calls were routinely discussed to ensure that care managers were adhering to this protocol.

After most calls, care managers would briefly communicate with referring physicians by e-mail. All outcome summary reports generated by the IVR outcome-monitoring process were sent to the referring physician as described above.

### DPC Workflow

The DPC process was intended to be as transparent as possible to the physicians. They were free to refer all, some, or none of their depressed patients at any point.

**Figure 1. The DPC care process: care manager case notes for patient D.F.**

<p>D.F. is a 42-year-old single woman. E-mail referral from Dr. X: Treated for depression by Dr. X for the last 3 years, otherwise healthy with no major comorbid conditions.</p> <p><b>Enrollment:</b> Intake call, initial PHQ-8 score of 15. Basic information given; encouraged to call with any questions. Information packet sent.</p> <p><b>3 Weeks after enrollment:</b> Follow-up care manager call, much better, PHQ = 7, no reported medication side effects.</p> <p><b>3 Months after enrollment:</b> Follow-up care manager call, patient worried that depression worse, PHQ = 21. Encouraged to schedule follow-up appointment with Dr. X. Seen 2 days later by Dr. X: increased fluoxetine. Unable to tolerate increased dose due to medication side effects. Phone consultation with psychiatric liaison, next steps determined.</p> <p><b>4.5 Months after enrollment:</b> Follow-up care manager call, PHQ = 19. Instructed to change antidepressant to venlafaxine per consultation plan.</p> <p><b>5.5 Months:</b> Follow-up care manager call, feeling much better, to call IVR to complete outcome summary.</p>
---

The referral protocol, customized to fit operating procedures at each site, worked in general as follows. The care manager spent time on-site at assigned practices to become familiar with faculty, staff, and practice workflow. Each physician at each site was given a list of his or her depression registry patients (prevalent cases) at the start of the intervention; at their discretion, physicians could refer any of these patients to the care manager for evaluation. New cases (incident cases) were identified through case-finding tools or by the physician during routine care and referred for evaluation and management. Educational materials regarding depression and the DPC program were posted in examination and waiting rooms at each site.

By design, enrollment was not limited to patients beginning acute-phase treatment for depression. The DPC project was designed to meet the needs of practicing physicians, who made it clear that their main need was for help with chronically ill, severely depressed, more-complex patients. Also by design, enrollment was not limited to patients meeting severity criteria for a major depressive episode. We realized that for some referred patients—particularly for those already receiving active treatment—a quantitative severity assessment (such as the PHQ-9) completed at the time of referral would yield scores that would not meet criteria for current major depressive disorder. We therefore expected that the DPC enrollment cohort would be heterogeneous and would include a high proportion of chronically depressed, treatment-resistant patients in various stages of active treatment. Although we postulated that this study cohort would bias the study against finding improved outcomes for the intervention, our intent was to assess the effectiveness of the DPC intervention for those patients who are the most difficult to manage in primary care.

Finally, participation was completely voluntary. Patients were free to opt out of the program at any time and reenter at their own or their clinician's discretion. The program was designed to provide long-term assessment and monitoring, so all enrollees were encouraged to continue in the program and complete infrequent (every 6 months) outcome monitoring calls even after they achieved recovery or remission or after active treatment ended.

### Clinical Setting of DPC Intervention

The intervention took place in 5 UMHS primary care sites: 2 single-specialty family medicine sites, and 3 multidisciplinary (general internal medicine, medicine-pediatrics, obstetrics-gynecology) sites. Five UMHS sites matched for size (physician patient panel size), composition (family medicine or multidisciplinary), and location were selected as usual care control sites. The

project accepted its first formal referral in May 2003, and sites were sequentially activated over several weeks.

### Methods of Evaluating Clinical Outcomes

The comprehensive mixed methods evaluation protocol for DPC program included quantitative clinical and economic analyses and qualitative analyses of physician, practice, and patient experiences with DPC project. Institutional review board approval for all aspects of the study was obtained from the University of Michigan Health System IRBMed. Data were standardized between the intervention (DPC) and usual care groups to enable a valid quasi-experimental comparison of clinical outcomes.

### Recruitment and Data Collection

Intervention patients were operationally defined as all patients referred to DPC who completed the intake telephone interview and at least 1 follow-up telephone contact with care managers. Data collected for this group included depression history and severity of current depression (at intake), as well as the PHQ-8, SF-12, and items on treatment adherence, functional impairment, and satisfaction with care (at each call). Data collection took place at the end of most care manager support calls, and most outcome data were collected by the IVR system at the end of each call. Later in the demonstration, a secure Web-based outcome assessment tool was available for enrollees to use. Both methods were chosen to minimize potential response bias. A small number of enrollees (mostly elderly) preferred to be asked outcome questions directly over the telephone by the care manager.

Usual care patients were recruited from UMHS control sites using the following procedure. At baseline, the UMHS depression registry was used to create a list of patients known to have a depressive disorder for each primary care clinician at each of the 5 intervention and 5 control sites. After receiving permission from the physicians, a letter signed by the patients' own physician was mailed to patients requesting their participation in a series of 4 mailed surveys. Patients who responded favorably to this request were approached by the project's research coordinator, who described the study and obtained written informed consent. To protect patient privacy, we were not allowed to contact potential participants directly and were unable to follow up with patients who did not respond to their personal clinician's request. Those who agreed to complete the series of questionnaires were sent a study packet at baseline, 6, 12, and 18 months. Survey items included depression history, current severity of depression, SF-12 and PHQ-8, and items assessing work disability, perceptions and beliefs about depression and its care,

and treatment adherence (full survey content available upon request). At each point, nonrespondents received follow-up by reminder postcards, a second mailing of packets, and telephone reminders following the method of Dillman.<sup>27</sup> A few patients successfully recruited to complete the series of questionnaires were also enrolled in the DPC intervention, creating a small cohort of DPC survey respondents. These patients were excluded from the usual care cohort in these analyses.

We undertook a structured electronic medical record review for both intervention and usual care groups to identify comorbid chronic health problems at baseline and to construct a structured case history of treatment, including specific treatment decisions, significant life events, and onset of new medical problems during the 18-month study period.

### Standardizing Outcome Assessment Intervals

To ensure a valid comparison of clinical outcomes, outcome assessment time points had to be matched between intervention and usual care patients. For DPC enrollees, the varying intervals between calls created a stream of irregularly spaced outcome assessments for each enrollee. To match the 6-month assessment intervals for the usual care group, we created an algorithm to identify whether a DPC outcome assessment call was completed within 6 weeks of the usual care 6-, 12-, and 18-month assessment intervals. Where available, data from these calls provided the 6-, 12-, and 18-month assessment points for intervention patients. Data from all other outcome assessment calls were not used in these comparisons.

### Controlling for DPC Dropouts

We hypothesized that dropout from the DPC program would not be random and carried out several analyses to assess whether the high rate of attrition in the DPC sample over time would bias our comparison of clinical outcomes. Comparisons of dropouts to retained cases at 12 and 18 months did not identify significant differences in demographic or clinical variables. Comparisons of last available remission rates and mean PHQ-8 scores between dropouts and retained cases at 12 and 18 months identified a consistent but not statistically significant pattern of higher remission rate and lower PHQ-8 scores in dropouts (data not shown). Analysis of missing data for individual intervention patients also showed that some dropouts returned to the program and the outcome-monitoring process when their symptoms recurred or worsened. These results suggest that dropout from the DPC program was not random; rather, it was more likely to occur for those with better outcomes. Those still active in the program at 12 and 18 months had relatively more severe symptoms.

We were also concerned about potential bias related to the low proportion of usual care patients who responded to the questionnaires, but we found no significant difference between survey respondents and nonrespondents on those demographic characteristics we could assess (age, sex, practice sites), and most respondents (85.7%) completed at least 3 of the 4 questionnaires and were included in all analyses. Based on this information, we believe that our usual care group was representative of the general population of chronically depressed primary care patients in our depression registry.

For analyses comparing changes in depression severity over 6, 12, and 18 months, we included all intervention and usual care patients with valid data for each time point. We also carried out a parallel analysis of changes in depression severity over 6, 12, and 18 months controlling for the duration of enrollment (a completer analysis). DPC program enrollees whose assessment time points provided an acceptable match for the 18-month time point plus at least 1 of the 6- and 12-month time points were included in the intervention sample for the completer analyses.

### Core Clinical Outcome Measures

For this analysis, the PHQ-8 score served as the primary outcome measure. The PHQ-8 was the well-validated PHQ-9 instrument,<sup>25,28-30</sup> omitting the ninth (self-harm) item, which we were legally constrained from asking early in the study. Possible scores on this instrument range from 0 (no depressive symptoms) to 24 (most severe); a change of 4 points is generally considered clinically meaningful. As for the full PHQ-9, remission at each time point was defined as a score of less than 5<sup>25</sup>; our reasoning was that patients at remission level would rarely score more than 0 on the self-harm item. A secondary outcome measure, the number of significantly limited-function work and activity days in past 30, was available for DPC enrollees only. The number of chronic medical problems (diabetes, coronary artery disease, congestive heart failure, asthma, chronic obstructive lung disease), health insurance type, sex, race/ethnicity, and history of depression were abstracted from the electronic medical record and used as covariates in multivariate analyses.

### Methods for Outcome Analyses

Core clinical outcomes for DPC enrollees and usual care patients (mean PHQ score, change from baseline PHQ score, and remission rate) were compared at 6, 12, and 18 months using  $\chi^2$  and *t* test methods as appropriate. Of 2,131 completed 8-item measures of the PHQ-8, missing data were noted on individual items in 34 (0.19%). Missing individual data points

on these measures were imputed as the mean of completed same-scale items to enable scoring for analysis. Confirmatory multivariate analysis was carried out using cumulative change in PHQ-9 score over the full 18-month period as the dependent variable in a linear regression model including all covariates. All analyses were conducted by members of the study team using statistical analysis software packages SPSS versions 16.0 and 17.0 (SPSS, Inc, Chicago, Illinois; <http://www.spss.com>) and Stata version 10.1 (StataCorp LP, College Station, Texas; <http://www.stata.com>).

## RESULTS

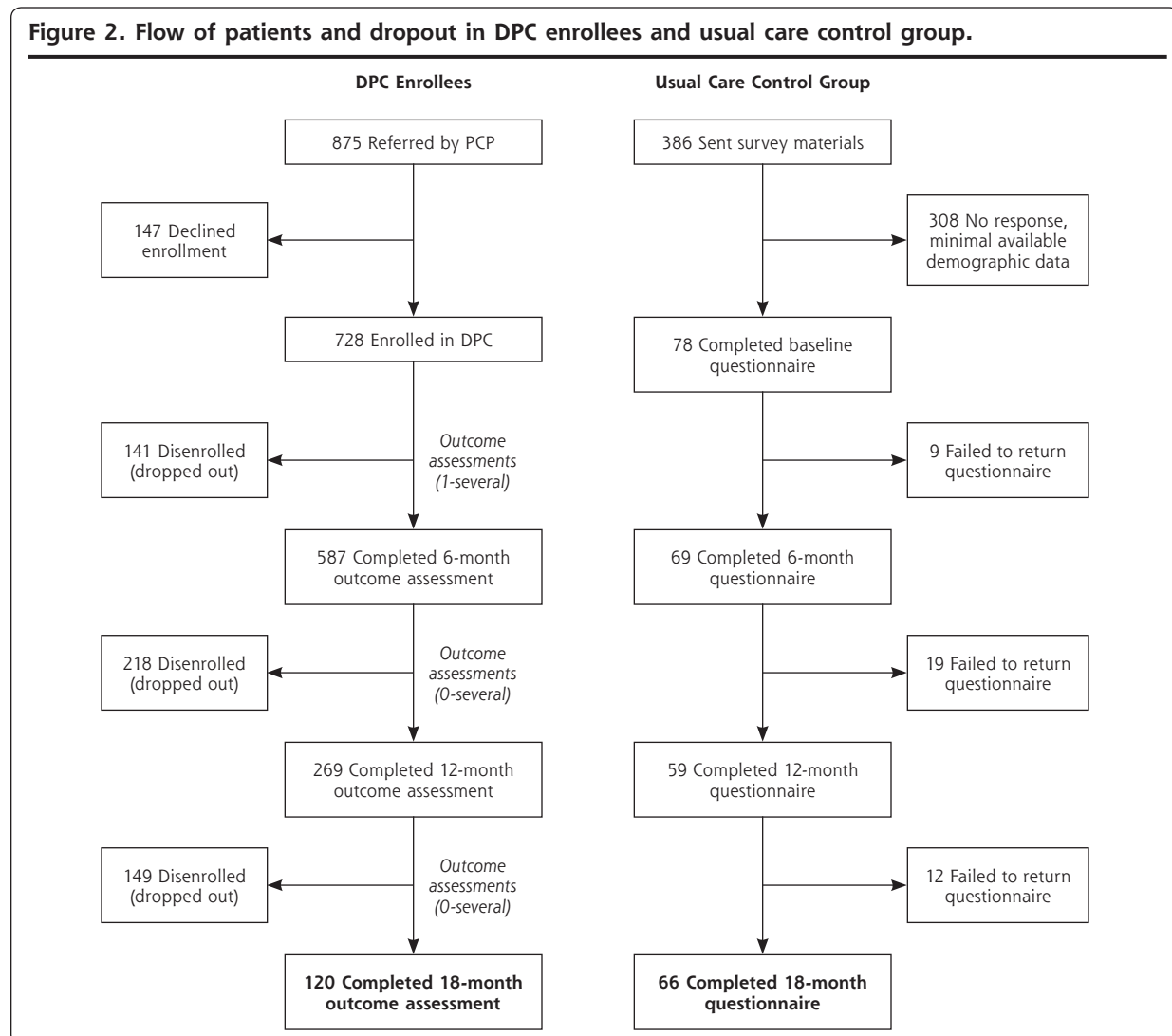
### Enrollment and Retention

Composition and accrual of intervention and control patients are displayed in Figure 2. As of January 1, 2005, 875 patients had been referred to the DPC

project, and 728 (83.2%) agreed to enroll in the program and completed the baseline assessment. Of the 728 enrollees, 587 completed the 6-month outcome assessment, 269 completed the 12-month outcome assessment, and 120 completed the 18-month outcome assessment. For the usual care group, 386 patients were sent the survey materials; 78 completed the baseline questionnaire, 69 completed the 6-month questionnaire, 59 completed the 12-month questionnaire, and 66 completed the 18-month questionnaire.

Referral rates to the DPC project were highest in sites where the mental health referral network was not well developed and lowest in the site where an active office-based depression monitoring program was introduced before the DPC project was initiated. Of those completing more than 1 outcome assessment call, 20.1% were assigned to level 1 (least complex cases), 69.3% assigned to level 2 (moderate complex-

**Figure 2. Flow of patients and dropout in DPC enrollees and usual care control group.**



ity), and 10.6% to level 3 (most complex, likely requiring specialty-level care).

Few level-1 patients were referred to the DPC program from sites 3 and 5. In site 3, a pilot study of point-of-care PHQ-9 monitoring for all patients whose problem list contained depression was in progress during the demonstration period, and site clinicians later stated that this monitoring protocol was sufficient in itself for management of level-1 patients (data not shown).

Across all 5 sites, clinician use of the program was highly variable. A few clinicians made no referrals to the DPC program, whereas most referred between 20 and 60 patients during the demonstration period. High-referring clinicians tended to refer patients entering new treatment episodes, as well as chronically depressed patients; others primarily referred complex patients with persistent symptoms despite treatment.

The mean number of outcome assessments per DPC enrollee was 4.0 (range 1-16), and the mean interval between the intake and initial follow-up monitoring call was 44.6 days (range 7-90 days); between calls 3 and 4 the mean interval was 67.7 days. The average length of each care manager telephone call as assessed by time-motion study was 18.1 minutes (range 0-50 minutes), with the average length of each follow-up monitoring call (obtained by excluding the longer intake calls from analysis) about 10 minutes. Each full-time equivalent care manager could handle a caseload of about 300 active enrollees. The caseload peaked at 350 enrollees per care manager, a sustainable load for short periods. The 875 patients referred to the DPC program represent approximately 40% of the total number of known depression registry patients at the outset of the demonstration, and the 728 enrolled represented almost 30% of the registry.

For the cohort of enrollees completing 18-month assessment, the mean number of outcome assessments was higher (about 6.0), but the average length of follow-up calls and interval between calls was about the same as for the full cohort. See the online Supplemental Appendix, available at <http://annfamned.org/cgi/content/full/8/5/387/DC1>,

for details. Supplemental Figure 1 and Supplemental Table 1 (available at <http://annfamned.org/cgi/content/full/8/5/387/DC1>) display the details of

the quarterly assessment of core outcomes for the full DPC enrollee population.

### Baseline Comparisons

There were no significant differences between DPC enrollees active in the program for 18 months and usual care patients in baseline demographic variables (Table 1). The mean age in both groups was about 46 years, more than two-thirds were female, and about 90% were covered by private managed-care or traditional health insurance plans. The mean PHQ-8 score at entry was 10.98 for the DPC enrollees and 9.94 for usual care patients. Usual care patients had slightly higher rates of comorbid medical problems, particularly diabetes (16.9% vs 10.8%), but this difference did not approach statistical significance. One in 3 patients had at least 1 comorbid medical problem.

The 2 groups were also similar in terms of history of depression and current depression treatment

**Table 1. Demographic Comparison of DPC Enrollees and Usual Care Patients Completing 18-Month Assessments**

Variable	DPC Enrollees (n = 120)	Usual Care (n = 65)	P Value
Mean age, y	46.1	46.0	.96 <sup>a</sup>
Female, %	79.2	70.7	.20 <sup>b</sup>
Insurance type, %			
Health maintenance organization	48.7	48.3	
Fee for service	39.3	40.3	
Medicare	9.4	6.5	
Other (Medicaid/safety net)	2.6	4.8	.79 <sup>b</sup>
Ethnic/racial minority, %	5.0	6.3	.72 <sup>b</sup>
Self-reported conditions, %			
Heart disease	7.5	4.6	.42 <sup>b</sup>
Diabetes	10.8	16.9	.24 <sup>b</sup>
Asthma	12.5	20.0	.17 <sup>b</sup>
Chronic obstructive pulmonary disease	5.0	6.2	.74 <sup>b</sup>
Comorbid medical problems, %			
0	68.3	61.5	
1	25.8	29.2	
2	4.2	7.7	
3 or more	1.7	1.5	.69 <sup>b</sup>
Known recurrent (>3 episodes) or chronic depression, %	65.8	76.6	.18 <sup>b</sup>
Taking antidepressant medication before intake/baseline, %	60.8	71.9	.18 <sup>b</sup>
Mean PHQ-8 score at entry	10.98	9.94	.24 <sup>a</sup>
Self-reported lifetime history, %	(n = 26)	(n = 65)	
Bipolar disorder <sup>c</sup>	15.4	15.6	.98 <sup>b</sup>
Anxiety disorder <sup>c</sup>	30.8	36.4	.61 <sup>b</sup>
Panic disorder <sup>c</sup>	19.2	9.2	.17 <sup>b</sup>

DPC = Depression in Primary Care; PHQ = Physician health Questionnaire.

<sup>a</sup> Assessed using t test.

<sup>b</sup> Assessed using  $\chi^2$ .

<sup>c</sup> Available for only the subset of DPC enrollees who completed questionnaires.

at baseline (Table 1). The proportion with known chronic or recurrent depression was high in both groups (65.8% vs 76.6%); although DPC enrollees were slightly less likely to be taking antidepressant medication at the time of their referral, this difference was not significant (60.8% vs 71.9%). Comparable data on mental health comorbidities were available for the small sample of DPC enrollees (n = 26) who completed the same series of 6-month surveys as usual care patients, and there were no differences between this small number of enrollees and usual care patients in the proportion with self-reported lifetime history of anxiety disorders or bipolar disorder.

### Outcome Comparisons

All primary outcome comparisons between DPC enrollees active in the program at each time point and usual care patients favored the DPC enrollees (Table 2). Mean PHQ-8 scores were lower for DPC enrollees than usual care patients at 6 months (7.48 vs 8.70), 12 months (6.87 vs 8.58), and 18 months (6.89 vs 8.77), but only the 12- and 18-month comparisons achieved statistical significance ( $P = .040$  and  $P = .019$ , respectively). The mean change in PHQ-8 score over each 6-month interval was more favorable for DPC enrollees than usual care patients, although only the baseline to 6-month comparison reached statistical significance ( $-4.40$  vs  $-1.59$ ,  $P < .001$ ). Over the full 18 months, DPC enrollees had a significantly greater reduction in mean PHQ-8 score than usual care patients ( $-4.17$  vs  $-1.07$ ,  $P < .001$ ). The proportion of DPC enrollees in remission was higher than usual care patients at 6 months (43.4%

vs 33.3%,  $P = .11$ ), 12 months (52.0% vs 33.9%,  $P = .012$ ), and 18 months (49.2% vs 27.3%,  $P = .004$ ). Repeating this set of comparisons including only the 120 DPC enrollees and 65 usual care patients for whom 18-month assessments were available yielded an almost identical pattern of results (see Supplemental Table 2, available online at <http://annfamned.org/cgi/content/full/8/5/387/DC1>, for details). The number-needed-to-treat to achieve remission at 18 months, compared with usual care, was 5 (95% CI, 2.8-12.6).

Confirmatory linear regression, using change in PHQ score over 18 months as dependent variable, confirmed that DPC enrollees had significantly more reduction in their depressive symptom burden than did usual care patients after controlling for age, sex, ethnicity, baseline severity, and comorbid medical illness (see Supplemental Table 3, available online at <http://annfamned.org/cgi/content/full/8/5/387/DC1>, for details).

### DISCUSSION

Our intent in this project was to assess the impact of the DPC intervention on the heterogeneous, chronically depressed patients commonly seen in primary care but rarely studied. We wanted to transform care management from its focus on short-term results for patients entering a new treatment episode to longer term results for patients receiving enhanced primary care treatment over time, and from intensive, structured short-term intervention protocols to more flexible, low-intensity and longer-term protocols that would

be sustainable in routine practice conditions. The subjects included in this analysis intentionally included a high proportion of chronically depressed, treatment-resistant patients who have other chronic health problems—a reasonable approximation of the real-world patients so difficult to treat successfully in primary care settings.

This study was not designed as a controlled clinical trial, and its limitations should be made clear. Clinicians and patients could not be randomly allocated to intervention or usual care, although practices were matched to the highest degree possible. The DPC intervention itself was heterogeneous. Clinicians and patients could make extensive use of the care manager for active collaborative care, or simply use the outcome summary report as a rough barometer of treatment response.

**Table 2. Clinical Outcome Comparison for DPC Enrollees and Usual Care Patients at 6, 12, and 18 Months**

Measure	DPC Enrollees		Usual Care		p Value
	n	Mean (SD)	n	Mean (SD)	
PHQ-8 score at entry	728	12.00 (5.65)	78	10.42 (5.64)	.190 <sup>a</sup>
PHQ-8 score at 6 mo	587	7.48 (5.58)	69	8.70 (5.16)	.084 <sup>a</sup>
PHQ-8 score at 12 mo	269	6.87 (5.88)	59	8.58 (5.24)	.040 <sup>a</sup>
PHQ-8 score at 18 mo	120	6.89 (5.32)	66	8.77 (4.85)	.019 <sup>a</sup>
Change in PHQ scores from baseline					
Baseline to 6 mo	584	-4.40 (5.71)	68	-1.59 (4.57)	<.001 <sup>a</sup>
6 to 12 mo	267	-0.42 (4.92)	58	0.04 (3.86)	.430 <sup>a</sup>
12 to 18 mo	113	0.03 (4.77)	57	0.36 (3.97)	.650 <sup>a</sup>
Baseline to 18 mo	119	-4.17 (6.37)	65	-1.07 (4.78)	<.001 <sup>a</sup>
Patients in remission		%		%	
At 6 mo	587	43.4	69	33.3	.110 <sup>b</sup>
At 12 mo	269	52.0	59	33.9	.012 <sup>b</sup>
At 18 mo	120	49.2	66	27.3	.004 <sup>b</sup>

DPC = Depression in Primary Care.

<sup>a</sup> Assessed using t test.

<sup>b</sup> Assessed using  $\chi^2$ ; with insurance type = managed care as comparator.



Patients could actively engage in the care management process and develop self-management skills, passively complete outcome assessments, or opt out of the program completely at any time. Some opted out, then returned to the program when symptoms returned. We could therefore not calculate a dose of the intervention for each patient.

More importantly, it was necessary to go through an extensive process to construct and validate roughly equivalent intervention and control groups and standardize data to carry out outcome comparisons. Although we believe that this process was successful, it was not the random allocation process used for controlled clinical trials. We could not establish major depressive disorder caseness for patients, as they were not referred at the beginning of a treatment episode, nor did we systematically assess for the presence of comorbid mental health problems at the outset of the demonstration. Our demographic data suggest, however, that most patients were chronically depressed and had substantial mental health comorbidity, and almost all had been included in the UMHS depression disease registry. Finally, we could not determine whether the usual care survey respondents were different in meaningful ways from nonrespondents or DPC enrollees, although they seemed similar in the limited ways we could assess. These limitations reflect the conscious tradeoff between experimental control and real-world generalizability, and our aim was to maximize the generalizability of our findings to routine primary care practice.

Despite these limitations, we believe that our core findings are robust. Our results confirm that this type of intervention is feasible and highly effective over time. The clinical protocol was successfully introduced in 5 primary care sites where change is constant, clinician turnover is high, and workflow is highly complex. We enrolled patients with chronic depressive symptoms despite (in many cases) ongoing treatment, a group that is of major clinical importance to primary care clinicians. Unlike acute-phase treatment studies, the relatively low PHQ-8 scores at baseline in this study do not represent minor or mild depression, but persistent symptoms despite active treatment. The program seemed highly acceptable to patients: 83% of those referred completed enrollment, a higher proportion than previously reported for these types of interventions,<sup>31</sup> and the great majority remained active in the program for at least 6 months. The intervention itself was efficient: the combination of flexible monitoring protocol with low-intensity care management (less than 10 minutes per follow-up call) meant that a full-time care manager could manage a caseload of more than 300 enrollees.

Most importantly, the intervention led to sustained improvement in core clinical outcomes over the longer term. Despite adverse patient selection incurred by the inclusion of chronically depressed and treatment-resistant patients, the disease-specific outcomes reported here are comparable to or better than published outcomes for interventions focused on acute-phase treatment. Almost one-half of patients (49.2%) met criteria for clinical remission at 18 months, almost doubling the remission rate seen in usual care delivered in our academic primary care setting (27.3%). The mean improvement in reduced-function days during the enrollees' time in the program was more than 36% across all sites. Every outcome comparison favored DPC enrollees, and the relative impact of the intervention was sustained at 18 months, long after the frequency and intensity of care management support had declined. The pattern of results found for the mean change in PHQ-8 score for each 6-month interval suggests that enrollees responded to the intervention during its more active intake and goal-setting phase, and they maintained those gains with less frequent contact over the longer term. We conclude that a low-intensity, tailored care management program based on the Chronic Care Model can lead to sustainable improvement in care for depression for chronically depressed patients found in real-world primary care practices.

All methods developed for this project were intended to be transportable to disease management programs for other chronic health conditions. By integrating care management tools and personnel across several related conditions (eg, depression, diabetes, and heart failure), it should be possible to achieve the scalability that will make integrated disease management feasible in the patient-centered medical home.<sup>32-35</sup> Further analyses of DPC results will compare costs and outcomes for subgroups of patients (for example, those with and without medical comorbidities), explore the direct relationships between intensity of care management and clinical outcomes, examine the impact of care management on clinician attitudes and practices, and patient beliefs, attitudes toward treatment, and adherence to treatment over time. The results of these analyses should provide valuable guidance to the development of chronic care management programs for the primary care setting.

**To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/cgi/content/full/8/5/387>.**

**Key words:** Depression; mental health; chronic disease; disease management; regression analysis

Submitted September 29, 2009; submitted, revised, March 12, 2010; accepted March 17, 2010.

**Funding support:** This demonstration project was supported by grants from the Robert Wood Johnson Foundation program, *Depression in Primary Care: Linking Clinical and Systems Strategies*.

## References

- Hirschfeld RM, Keller MB, Panico S, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *JAMA*. 1997;277(4):333-340.
- Hunkeler EM, Meresman JF, Hargreaves WA, et al. Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Arch Fam Med*. 2000;9(8):700-708.
- Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA*. 2000;283(2):212-220.
- Unützer J, Rubenstein L, Katon WJ, et al. Two-year effects of quality improvement programs on medication management for depression. *Arch Gen Psychiatry*. 2001;58(10):935-942.
- Unützer J, Katon W, Callahan CM, et al.; IMPACT Investigators. Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA*. 2002;288(22):2836-2845.
- Simon GE, Ludman EJ, Tutty S, Operskalski B, Korff MV. Telephone psychotherapy and telephone care management for primary care patients starting antidepressant treatment: a randomized controlled trial. *JAMA*. 2004;292(8):935-942.
- Dietrich AJ, Oxman TE, Williams JW Jr, et al. Re-engineering systems for the treatment of depression in primary care: cluster randomized controlled trial. *BMJ*. 2004;329(7466):602.
- Neumeier-Gromen A, Lampert T, Stark K, Kallischnigg G. Disease management programs for depression: a systematic review and meta-analysis of randomized controlled trials. *Med Care*. 2004;42(12):1211-1221.
- Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care: a systematic review. *JAMA*. 2003;289(23):3145-3151.
- Katon WJ, Unützer J, Simon G. Treatment of depression in primary care: where we are, where we can go. *Med Care*. 2004;42(12):1153-1157.
- Lin EH, Simon GE, Katon WJ, et al. Can enhanced acute-phase treatment of depression improve long-term outcomes? A report of randomized trials in primary care. *Am J Psychiatry*. 1999;156(4):643-645.
- Hunkeler EM, Katon W, Tang L, et al. Long term outcomes from the IMPACT randomized trial for depressed elderly patients in primary care. *BMJ*. 2006;332(7536):259-263.
- Wells K, Sherbourne C, Schoenbaum M, et al. Five-year impact of quality improvement for depression: results of a group-level randomized controlled trial. *Arch Gen Psychiatry*. 2004;61(4):378-386.
- Sherbourne CD, Wells KB, Duan N, et al. Long-term effectiveness of disseminating quality improvement for depression in primary care. *Arch Gen Psychiatry*. 2001;58(7):696-703.
- Kilbourne AM, Schulberg HC, Post EP, Rollman BL, Belnap BH, Pincus HA. Translating evidence-based depression management services to community-based primary care practices. *Milbank Q*. 2004;82(4):631-659.
- Kessler RC, Stang PE, Wittchen HU, Ustun TB, Roy-Burne PP, Walters EE. Lifetime panic-depression comorbidity in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1998;55(9):801-808.
- Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617-627.
- Olfson M, Shea S, Feder A, et al. Prevalence of anxiety, depression, and substance use disorders in an urban general medicine practice. *Arch Fam Med*. 2000;9(9):876-883.
- Zimmerman M, McDermet W, Mattia JI. Frequency of anxiety disorders in psychiatric outpatients with major depressive disorder. *Am J Psychiatry*. 2000;157(8):1337-1340.
- Schwenk TL, Evans DL, Laden SK, Lewis L. Treatment outcome and physician-patient communication in primary care patients with chronic, recurrent depression. *Am J Psychiatry*. 2004;161(10):1892-1901.
- Pincus HA, Pechura C, Keyser D, Bachman J, Houtsinger JK. Depression in primary care: learning lessons in a national quality improvement program. *Adm Policy Ment Health*. 2006;33(1):2-15.
- Pincus HA, Pechura CM, Elinson L, Pettit AR. Depression in primary care: linking clinical and systems strategies. *Gen Hosp Psychiatry*. 2001;23(6):311-318.
- Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. *JAMA*. 2002;288(14):1775-1779.
- Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, part 2. *JAMA*. 2002;288(15):1909-1914.
- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
- Ware JE Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233.
- Dillman DA. *Mail and Internet Surveys: The Tailored Design Method*. New York, NY: John Wiley & Sons; 2000.
- Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*. 2004;81(1):61-66.
- Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med*. 2006;21(6):547-552.
- Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Med Care*. 2004;42(12):1194-1201.
- Datto CJ, Thompson R, Horowitz D, Disbot M, Oslin DW. The pilot study of a telephone disease management program for depression. *Gen Hosp Psychiatry*. 2003;25(3):169-177.
- Miller BF, Mendenhall TJ, Malik AD. Integrated primary care: an inclusive three-world view through process metrics and empirical discrimination. *J Clin Psychol Med Settings*. 2009;16(1):21-30.
- Kessler R, Stafford D, Messier R. The problem of integrating behavioral health in the medical home and the questions it leads to. *J Clin Psychol Med Settings*. 2009;16(1):4-12.
- Nutting PA, Gallagher KM, Riley K, White S, Dietrich AJ, Dickinson WP. Implementing a depression improvement intervention in five health care organizations: experience from the RESPECT-Depression trial. *Adm Policy Ment Health*. 2007;34(2):127-137.
- Nutting PA, Miller WL, Crabtree BF, Jaen CR, Stewart EE, Stange KC. Initial lessons from the first national demonstration project on practice transformation to a patient-centered medical home. *Ann Fam Med*. 2009;7(3):254-260.